

# Convergence of Depth from Texture and Depth from Disparity in Macaque Inferior Temporal Cortex

Yan Liu, Rufin Vogels, and Guy A. Orban

Laboratorium voor Neuro-en Psychofysiologie, Katholieke Universiteit Leuven Medical School, Campus Gasthuisberg, B3000 Leuven, Belgium

The visual system must reconstruct the three-dimensional structure of an object from two-dimensional retinal images. Previous research has shown that macaque inferior temporal (IT) neurons, although belonging to the ventral visual stream, code for depth defined by binocular disparity gradients. Here, we demonstrate that macaque IT neurons also code for depth defined by texture gradients, a monocular depth cue. Single IT neurons were selective for the tilt of texture-defined surfaces, and the tilt preferences of individual neurons remained the same, whether surfaces were defined by texture or disparity cues. Furthermore, the tilt preference was invariant over different types of textures and slants, suggesting an abstract representation of surface tilt in ventral visual cortex.

**Key words:** depth cues; inferior temporal; macaque; texture; disparity; extrastriate

## Introduction

The primate brain computes depth structure using diverse visual cues such as binocular disparity, texture gradients, and shading (Howard and Rogers, 2002). Recently, a series of single-cell studies in macaques have shown tunings for gradients in binocular disparity (stereo) in dorsal (Taira et al., 2000; Nguyenkim and DeAngelis, 2003) as well as ventral visual areas (Janssen et al., 1999, 2000a, 2001; Hinkle and Connor, 2002). However, much less is known about the coding of texture gradients, another potent depth cue (Gibson, 1950), as nicely demonstrated in the paintings of Vasarely. Only very recently, Tsutsui et al. (2002) showed that neurons in the dorsal visual stream area CIP are tuned for the texture-defined three-dimensional (3D) orientation of planar surfaces. A subset of these parietal neurons was 3D orientation selective for surfaces defined by disparity or texture, indicating a convergence of the two-depth cues at the single-cell level.

We have shown previously (Janssen et al., 1999, 2000a,b, 2001) that a region in the macaque inferior temporal (IT) cortex, part of the ventral visual stream, contains neurons selective for curvature and gradients defined by disparity, and we therefore suggested that these neurons code for the depth structure of objects. In the present study, we address the question whether these IT neurons code for texture-defined 3D shape. A number of IT neurons has been shown to be selective for texture (Tanaka et al., 1991; Komatsu and Ideura, 1993), but it is unknown to what degree the texture sensitivity relates to 3D shape or to two-dimensional (2D) texture pattern differences (as in materials). Indeed, selectivity for texture gradients does not necessarily imply that the neurons code for 3D orientation, because the latter stimulus feature covaries with variations in 2D texture

elements. To determine whether the texture gradient selectivity relates to 3D or to 2D coding, we measured the responses of single IT neurons to both texture and disparity gradients. If the 3D orientation preferences for texture-defined stimuli correlate with those defined by disparity, this would indicate that these neurons code for 3D texture.

To facilitate a comparison of the present results in the ventral visual stream with those in the dorsal stream (Tsutsui et al., 2002), we examined the selectivity of IT neurons using planar instead of curved surfaces (Janssen et al., 1999, 2000a,b, 2001). In addition to comparing the tilt tunings for disparity-defined and texture-defined gradients, we determined whether the tilt tuning is invariant with respect to various parameters (i.e., texture type and slant) and to the presentation mode (monocular versus binocular) of the texture. Such invariance would facilitate the computation of surface tilt.

## Materials and Methods

**Stimuli.** The planar textures were perspective projections of a dot texture, line texture, combination of the latter two, and of one with square texels. These were rendered at four tilts in 90° steps with a 45° slant (see Fig. 1) in the main test. The rendering was performed using commercial software (3DS max 4; Autodesk, Montreal, Canada) using a wide camera angle that produced a salient perception of depth for the relatively small stimulus size used in this study. Random dot stereograms (50% density) depicting planar (disparity range, 0.80°) were produced as those by Janssen et al. (2001). These surfaces were centered on zero disparity. We did not attempt to equate the slant of the texture- and disparity-defined gradients, the latter depending on the disparity range for a given stimulus size. The border of each surface or its apparent contour could be any of six simple shapes. Four of these apparent contours were taken from the 2D shapes used by Janssen et al. (2001). The other two apparent contours were a square as for the images in Figure 1 (top) and a circle (bottom). The apparent contour of the stereo-defined surfaces was set at zero disparity. The stimuli were ~7° in size.

The stimulus presentation apparatus was identical to that of our previous studies (Janssen et al., 2000a,b, 2001, 2003). The stimuli were presented dichoptically (or monocularly) by means of a double pair of ferroelectric liquid crystal shutters operating at 60 Hz (Janssen et al., 1999, 2000a) placed in front of the monkey's eyes. The stimuli were presented

Received Jan. 15, 2004; revised March 5, 2004; accepted March 8, 2004.

This work was supported by Geneeskundige Stichting Koningin Elizabeth, Fonds voor Wetenschappelijk Onderzoek-Vlaanderen, Insight2+, Grant Geconcerteerde Onderzoeksactie 2000/11. We gratefully acknowledge the technical help from M. De Paep, P. Kayenbergh, G. Meulemans, and G. Vanparijs and the critical reading by Dr. Steve Raiguel.

Correspondence should be addressed to Rufin Vogels, Laboratorium voor Neuro- en Psychofysiologie, Katholieke Universiteit Leuven Medical School, Campus Gasthuisberg, B3000 Leuven, Belgium. E-mail: rufin.vogels@med.kuleuven.ac.be.  
DOI:10.1523/JNEUROSCI.0150-04.2004

Copyright © 2004 Society for Neuroscience 0270-6474/04/243795-06\$15.00/0

foveally during steady fixation of a 0.19° target that was superimposed on the stimulus. The viewing distance was 86 cm.

**Single-cell recording.** Standard extracellular single-cell recordings were made in the rostral lower bank of the superior temporal sulcus (STS) in each hemisphere of two rhesus monkeys (M1 and M2; *Macaca mulatta*) using previously published procedures (Janssen et al., 2000a,b, 2001, 2003). Both subjects showed stereopsis as demonstrated by psychophysical testing of the discrimination of disparity-defined curvature. A head post, a scleral search coil, and a recording well were implanted under sterile conditions and deep isoflurane anesthesia. Horizontal and vertical movements of one eye were recorded with the scleral search coil technique (Judge et al., 1980) at 200 Hz.

As described in previous reports (Janssen et al., 2000a,b, 2001, 2003), we used a vertical approach to the rostral lower bank of the STS. Before the present study, we recorded the responses to random dot stereograms of curved surfaces in one hemisphere in each of the two animals. In those recordings, we identified a region in the lower bank of the STS that responded selectively to the curvature (convex versus concave) of the stereo-defined 3D shapes. The present recordings in these two hemispheres were performed at the same or similar guiding tube positions and at depths similar to those in the studies using the curved surfaces. The responsive region of the STS in the other hemisphere of each of the animals was determined using the disparity- and texture-defined planar surfaces. The localization of the recording site as being the rostral lower bank of the STS is based on the transitions of gray and white matter observed during the recordings as well as on structural magnetic resonance imaging (MRI) and computed tomography (CT) scanning with a guiding tube *in situ* as described by Janssen et al. (2000b). The animal was placed in a plastic stereotaxic frame during the MRI and CT scans. All animal procedures were approved by the K. U. Leuven Ethical Committee.

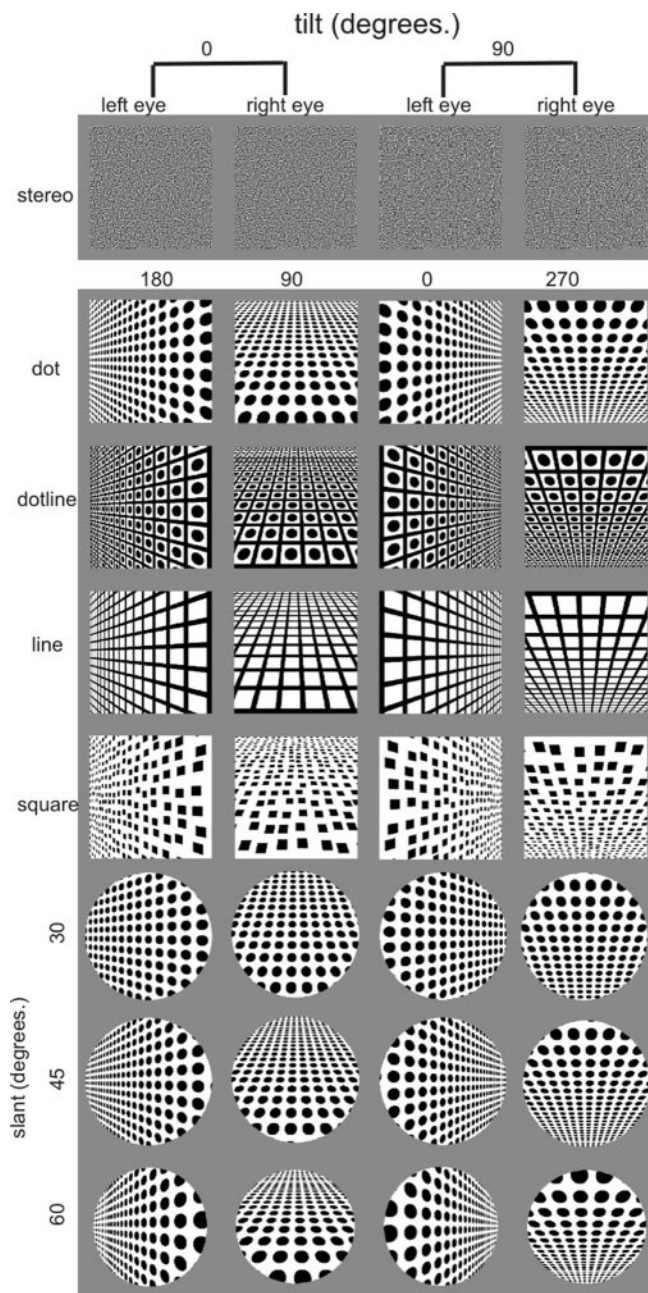
The monkeys performed a fixation task during recording. After 1000 msec of stable fixation, the stimulus was presented for 600 msec. The monkeys were rewarded with a drop of apple juice for maintaining fixation during the entire duration of the trial. The square fixation windows measured 1.8 and 3° in M1 and M2, respectively.

Responsive neurons were sought using the four tilts of the dot textures and disparity gradients (see Fig. 1). In each recording session, we searched for responsive neurons using three apparent contours that had been randomly chosen from a total of six. Thus, the search test consisted of 24 conditions (four tilts times two depth cues times three apparent contours). Responsive neurons were tested using the preferred apparent contour. The main test consisted of interleaved binocular presentations of the four tilts of the disparity-defined and dot texture-defined planes and monocular presentations of the disparity conditions. For 41 neurons, this test included monocular presentations of the dot texture gradients. In 62 neurons, the three other texture types were presented binocularly, interleaved with the dot texture and stereograms. This test was followed by a control comparing three different slants of the dot-texture pattern.

**Data analysis.** Responses were computed by subtracting the number of spikes counted in a 400 msec interval immediately preceding stimulus onset (baseline window) from the number of spikes in a 400 msec interval starting 80 msec after stimulus onset (response window) (Janssen et al., 2000a,b, 2001, 2003). The significance of responses and tilt selectivity was assessed using an ANOVA with tilt as a between-trials factor and the number of spikes in the baseline and response window as a within-trials factor. A significant main effect of the latter response factor ( $p < 0.05$ ) or a significant interaction of the two factors ( $p < 0.05$ ) designated a significant response for a given neuron. A significant interaction ( $p < 0.05$ ) or main effect of tilt ( $p < 0.05$ ) defined significant tilt selectivity.

To compute the preferred tilt and degree of selectivity, the mean responses for each of the four tilts were converted into vectors so that the vector angle corresponded to surface tilt and the vector length to the response strength. The preferred tilt was defined as the direction of the sum vector of these vectors. The selectivity index (SI) (Vogels and Orban, 1994) is the normalized length of this sum vector of the responses and can vary between 0 (equal responses to the different tilts) and 1 (response to one tilt only):

$$SI = [(\sum [R_i \cdot \sin(T_i)])^2 + (\sum [R_i \cdot \cos(T_i)])^2]^{1/2} / \sum R_i$$



**Figure 1.** Stimuli. Top, Two examples of a pair of random dot patterns that, when presented dichoptically, produce disparity gradients of the indicated tilts (top row) and illustrations of the texture gradients of four different tilts (columns; tilt indicated on top) of each texture type (bottom rows). The apparent contour is a square. Bottom, Dot texture pattern projected at four different tilts (columns) and three different slants (rows) with a circular contour.

with  $R_i$  being the response of the neuron at tilt  $T_i$ .

A second index of tilt selectivity that we used was the tilt modulation index (TMI) (Nguyenkim and DeAngelis, 2003). The latter is defined as the difference between the response to the best tilt and the response to the worst tilt, divided by the response to the best tilt. Other analyses are described in Results.

## Results

Responsive neurons were sought in the lower bank of the rostral part of the STS (TEs) (Janssen et al., 1999, 2000a,b, 2001) while presenting images of dot texture-defined surfaces and of random dot patterns containing disparity gradients. For each gradient type, four different tilts in 90° steps (Fig. 1) combined with three

**Table 1. Responses of IT neurons ( $n = 115$ ) to texture and disparity gradients**

	Responsive cells ( $n$ )	Mean SI	Mean TMI
Dot texture gradient	105	0.30	0.81
Disparity gradient	95	0.33	0.90

of a total of six apparent contours were used as search stimuli in each session. Neurons responsive to at least one of the stimuli were formally tested by presenting texture- and disparity-defined surfaces at each of the four tilts for the preferred apparent contour.

### Responses and selectivity for tilt defined by dot texture and disparity gradients

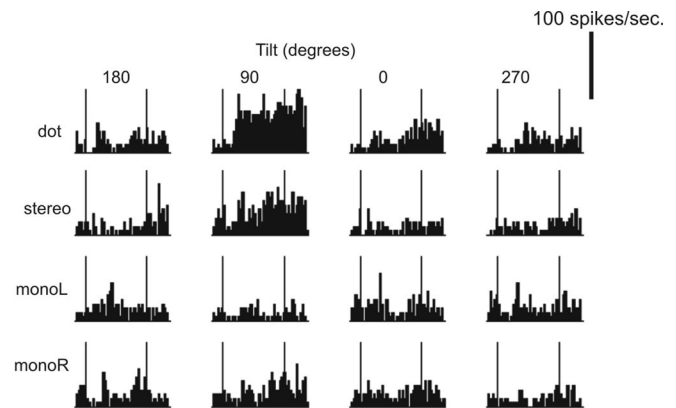
Of 115 (M1, 79; M2, 36) IT neurons that were judged to be responsive to at least one of the search stimuli, 105 (91%) and 95 (83%) responded significantly to the dot texture and disparity gradients, respectively. On the basis of CT scans, the Horsley–Clark anteroposterior and mediolateral coordinates of the selective neurons ranged from 15 to 19 mm and from 19 to 22 mm, respectively. Neurons selective for the texture and disparity gradients were found for each of the guiding tube positions.

On average, the responses to the dot gradients (mean best response, 21.1 spikes/sec) exceeded those for the random dot stimuli (16.5 spikes/sec; Wilcoxon matched pair test;  $p < 0.0002$ ). A majority of the responsive neurons showed a significant selectivity (ANOVA;  $p < 0.05$ ) for the tilt of the texture gradients (74 of 105 neurons; 70%) and the disparity gradients (65 of 95; 68%). The degree of tilt selectivity, as assessed by two separate indices (SI and TMI; see Materials and Methods), was similar for the texture- and disparity-defined stimuli (Table 1).

To confirm that the selectivity for the disparity gradient images is actually attributable to disparity rather than incidental differences in the random dot patterns constituting the different tilts, we tested each neuron with monocular presentations of the left and right eye disparity stimuli. Only five neurons showed significant effects of tilt for these monocular controls, indicating that the tilt tuning was attributable to genuine disparity selectivity in the large majority of neurons.

A large proportion (86 of 115; 75%) of the neurons responded to both the texture and disparity gradients, and 90% (77 of 86) of those showed tilt selectivity for the texture or the disparity cue. Approximately half (41 of 77; 53%) of these selective neurons exhibited significant tilt selectivity both for gradients defined by disparity and by texture (M1, 27 neurons; M2, 14 neurons). If the tuning for texture gradients actually reflect tuning for 3D tilt rather than responses to the 2D textures, one would expect that the tilt preference of these 41 neurons would be similar for the two depth cues. Figure 2 shows a neuron for which this was indeed the case: this neuron responded selectively for tilt, with similar preferred tilts for the two cues. Also note that the tilt selectivity was absent when the left and right images of the stereo stimuli were presented alone (monocular controls).

To relate 3D tilt tuning for the two cues at the population level, we aligned the tuning curve of each neuron to its preferred tilt in the disparity cue conditions. The tuning curves of the texture and monocular cue conditions were then rotated by this same angle for a given neuron. The average tuning curves for the 41 neurons that were tilt selective for the two depth cues (Fig. 3C) demonstrate that the mean tilt tuning for the two cues is very similar. This is reflected in the high correlation ( $r = 0.81$ ;  $p < 0.005$ ;  $n = 41$ ) between the preferred tilts for the disparity and the texture cue (Fig. 3A). The distribution of the difference of preferred tilt



**Figure 2.** Responses of a single inferior temporal neuron to binocular presentations of the dot textures (top row) and stereo stimuli (second row) portraying four tilts (columns). The two bottom rows show responses to monocular presentations of the left (MonoL) and right (MonoR) eye stereo images. The onset and offset of the stimulus (600 msec duration) is indicated by the vertical bars.

between the two conditions differed significantly from a uniform distribution ( $\chi^2 = 36.1$ ;  $p < 0.00001$ ) (Fig. 3B). In fact, in half of these 41 neurons, the preferred tilts for the two cues differed by  $<40^\circ$  and exceeded  $90^\circ$  for only 15% of the neurons (Fig. 3B). Such a high degree of correspondence would be unexpected if the responses to the texture gradients merely reflected 2D texture instead of 3D selectivity. Therefore, we conclude that single IT neurons code for depth from texture. The high degree of correlation between preferred tilts also excludes the possibility that the selectivity for disparity gradients is attributable to improbable tilt-dependent changes in vergence angle or receptive field inhomogeneities instead of a genuine tilt tuning.

The distribution of the preferred tilts defined by texture was not uniformly distributed in the sample of neurons that was selective for the disparity and texture gradients ( $\chi^2 = 9.2$ ;  $p < 0.05$ ), with most neurons concentrated between  $0$  and  $270^\circ$  of tilt. A similar but not significant trend ( $\chi^2 = 3.7$ ; NS) was present for the preferred tilts defined by disparity. However, the observed nonuniform distribution of preferred tilts might reflect a sampling bias rather than a genuine anisotropy.

### Time course of tilt tuning for the two cues compared

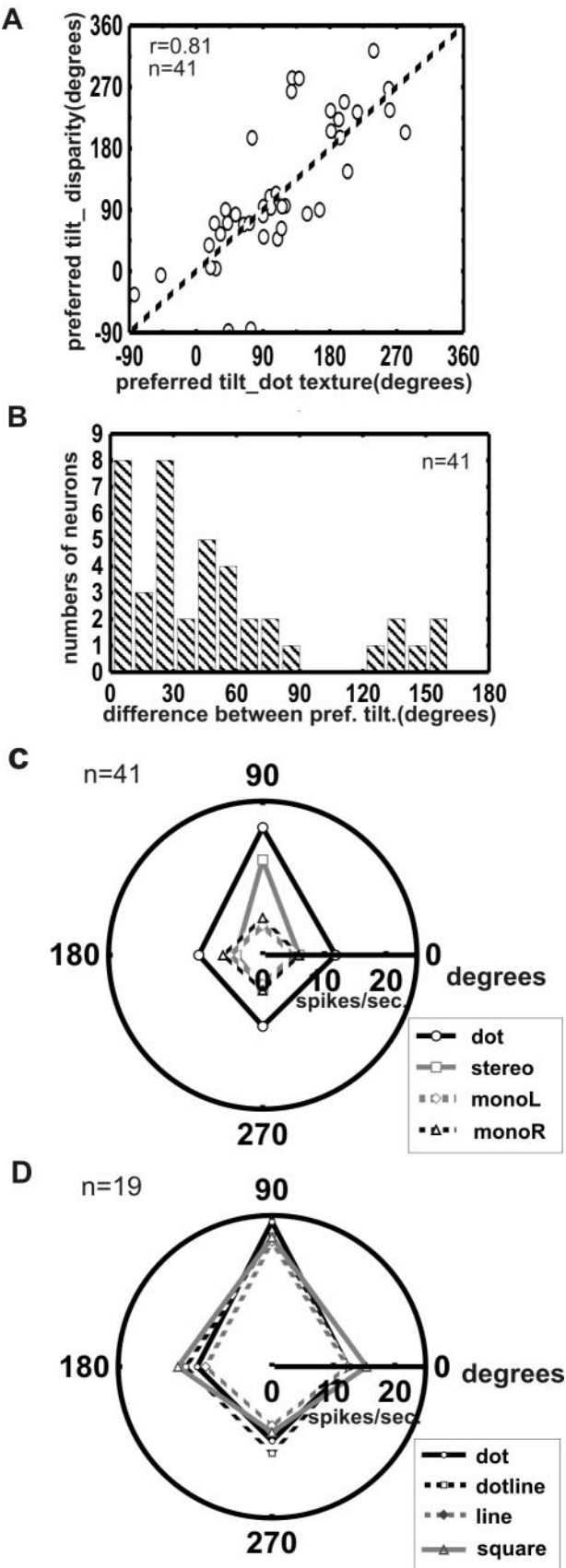
The time courses of the tilt tuning were examined for the 41 neurons that were tilt selective for both cues. For each cue and neuron, the best and worst tilts were determined, and the normalized peristimulus time histograms of the different neurons were averaged for each condition. Figure 4A shows that the population response was greater for the texture than for the stereo random dot patterns, in agreement with the previous analysis. Interestingly, the population responses for the two cues had similar time courses. Subtracting the population responses for best and worst tilt (Fig. 4B) indicated that the time course of the tilt selectivity was also similar for the two cues.

### Invariance of tilt tuning to texture type

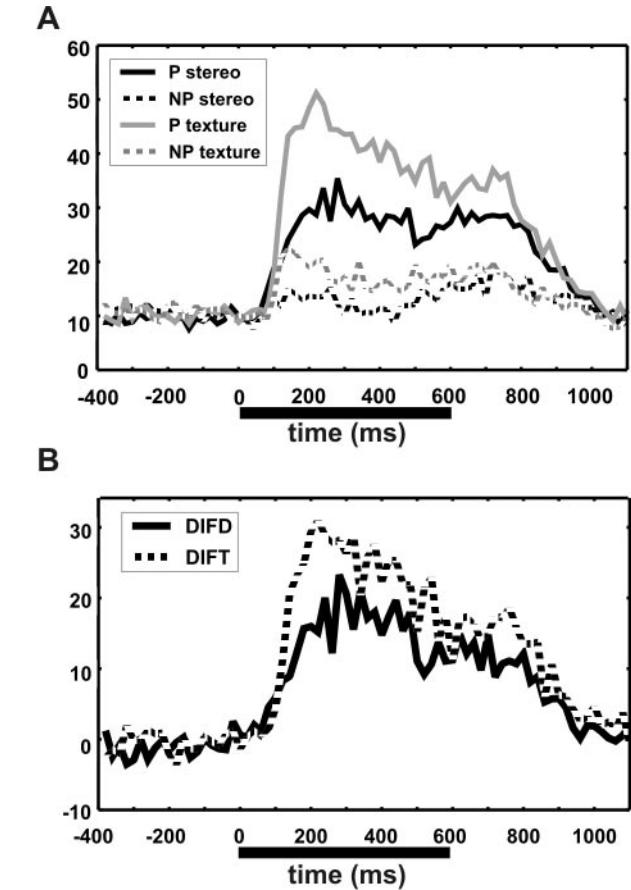
To determine whether the tilt tuning was retained using different texture patterns, we tested 62 neurons with four different types of textures (Fig. 1). Note that the position of the square texels was randomized in the fourth texture pattern, making it quite different from the others.

Overall, the mean response strengths and degrees of tilt selectivity were similar for the four textures (Table 2). ANOVAs with





**Figure 3.** The tuning of inferior temporal neurons to tilt portrayed by different textures and disparity gradients is compared. *A*, Correlation of the preferred tilt for the dot texture and disparity gradients for those neurons selective for both cues. *B*, Absolute difference in preferred tilt for the dot texture and disparity gradients for the same neurons in *A*. *C*, Population tuning; polar plot of the mean



**Figure 4.** Population analysis of time courses of responses to the disparity- and dot texture-defined surfaces for 41 neurons tilt selective for both depth cues. *A*, Population response histograms for the best, preferred, texture-defined (P texture), and disparity-defined tilt (P stereo) and for the worst, nonpreferred, texture-defined (NP texture), and disparity-defined tilt (NP stereo). Before averaging, the maximum of four peristimulus time histograms was normalized to 100% for each neuron (bin width, 20 msec). The black bar indicates stimulus presentation. *B*, Differences between the population responses to the best and worst tilt in the texture (DIFT) and disparity (DIFD) conditions. Same neurons and conventions as in *A*.

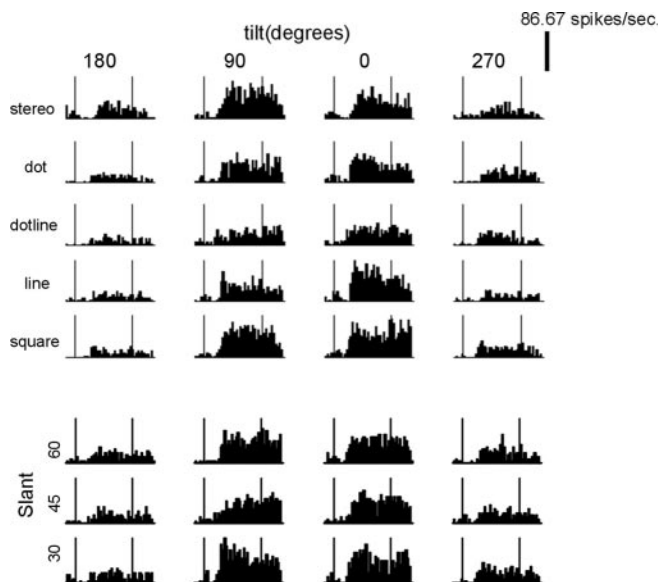
texture type and tilt as factors showed that the response of most neurons (51 of 62; 82%) depended significantly on the texture type. Nonetheless, 52 of 62 (84%) of the neurons gave significant responses for all four texture types. The tilt selectivity was significant for all four texture types in only 19 of 52 (36%) neurons, indicating that individual neurons vary in their degree of selectivity for the different texture types.

Figure 5 shows one neuron that was tilt selective for all four texture types. Although the response strength varied with texture type, the tilt tuning remained relatively constant. Similarity of tilt preferences across texture types was generally observed in those neurons selective for each of the four texture types, as demonstrated by plotting the average tuning curves using the dot texture as reference (Fig. 3*D*). Indeed, the correlations of the preferred tilt for each of the six pairs of textures ranged from 0.87 to 0.91 ( $n = 19$ ) (Table 3), all highly significant. Also, the average preferred tilt tested with the disparity gradients closely matched the

response of the neurons selective for both cues. The preferred tilt in the stereo condition of each neuron was aligned with 90°. MonoL and MonoR, Mean responses to left and right eye stereo images, respectively. *D*, Population tuning; polar plot of the mean response of neurons selective for each of the four texture types. The preferred tilt in the dot texture condition was aligned with 90°.

**Table 2. Responses of IT neurons ( $n = 62$ ) to different types of textures**

	Mean best response	Responsive cells ( $n$ )	Mean SI	Mean TMI
Dot texture	23.0 spikes/sec	58	0.26	0.75
Line-dot texture	23.8 spikes/sec	56	0.28	0.74
Line texture	22.2 spikes/sec	58	0.31	0.87
Square texture	22.6 spikes/sec	56	0.24	0.88

**Figure 5.** Responses of a single inferior temporal neuron to binocular presentations of the texture gradient and disparity gradients portraying four tilts (columns). The top panel shows the response of the neuron to stereo images and four different texture types. The bottom panel shows the responses of the same neuron to dot textures portrayed with four different tilts and three slants. Conventions are the same as in Figure 2.**Table 3. Correlation of preferred tilt among texture types for tilt-selective neurons**

Texture types	Neurons selective for all four texture types	Neurons not selective for all four texture types
Dot and line-dot	$r = 0.91$ ( $n = 19$ )	$r = 0.95$ ( $n = 11^*$ )
Dot and line	$r = 0.89$ ( $n = 19$ )	$r = 0.94$ ( $n = 7^*$ )
Dot and square	$r = 0.90$ ( $n = 19$ )	$r = 0.83$ ( $n = 6^*$ )
Line-dot and square	$r = 0.87$ ( $n = 19$ )	$r = 0.80$ ( $n = 13^*$ )
Line-dot and line	$r = 0.90$ ( $n = 19$ )	$r = 0.86$ ( $n = 11^*$ )
Line and square	$r = 0.91$ ( $n = 19$ )	$r = 0.92$ ( $n = 9^*$ )

\*Number of neurons selective for the two texture types listed in the first column. All correlations were statistically significant (\* $p < 0.05$ ).

preferred tilt of the four texture patterns (Fig. 5), again indicating that the texture tuning is related to 3D coding.

The invariance of tilt preference with respect to texture type was also present in those neurons that responded to all four texture types but were selective for at least two but less than four texture types. For each pair of texture types, we selected those neurons (sample size ranging from 6 to 13) (Table 3) that responded selectively to the two texture types, excluding the 19 neurons that were selective for each of the four textures. For these selected samples of neurons, the correlations of the preferred tilt for each of the six pairs of textures were at least 0.80, all of which were statistically significant (Table 3). Thus, the invariance of tilt preference to texture type is not an exclusive property of the 19 neurons that were selective for each texture type but is partially shared by other neurons.

### Invariance of tilt tuning to slant

To examine any possible effect of slant on tilt tuning, we tested 41 responsive neurons with the dot texture patterns using slants of 30, 45, and 60° combined with the four tilts (Fig. 1). A two-way ANOVA showed a main effect of slant and tilt in 14 (34%) and 34 (78%) of the neurons, respectively. More than half of the neurons (8 of 14) tuned for slant preferred the largest slant (60°). Thirty four percent (14) of the neurons showed a significant interaction between tilt and slant and, in 9 of 14 neurons, this was because of the absence of tilt selectivity for one of the three slants. Importantly, the preferred tilt of the neurons was mostly independent of slant, as the example neuron in Figure 5 demonstrates. For the 34 neurons that showed a significant tilt tuning, the correlations of the preferred tilt for the three possible pairings of the different slants ranged between 0.68 and 0.86 ( $n = 34$ ; all  $p < 0.05$ ).

### Invariance of tilt tuning to binocular versus monocular presentations

The binocular presentation of texture gradients implies a conflict between the texture cue and disparity cue of that stimulus, because the latter signals a flat surface. Despite this cue conflict, human and monkey observers (Tsutsui et al., 2002) perceived the surface tilt in the binocular presentations. For monocular presentations, the disparity cue is absent, thus removing the conflict between the two cues. In 41 neurons that were responsive to at least one of the presentation modes, we compared monocular and binocular presentations of the same textures. A two-way ANOVA with presentation mode and tilt as factors showed a significant effect of the former factor in 34% of the neurons and an interaction in 15%. The majority (68%; 28) of these neurons were selective for tilt (main tilt effect or interaction) and, as expected, the preferred tilt for the two presentation modes correlated significantly ( $r = 0.88$ ;  $n = 28$ ). Interestingly, the average degree of tilt selectivity was significantly larger for the monocular (SI,  $0.32 \pm 0.17$ ) than for the binocular presentations (SI,  $0.26 \pm 0.14$ ; Wilcoxon matched pair sample test;  $p < 0.02$ ;  $n = 40$  neurons responsive in both presentation modes).

### Discussion

We showed that a population of IT neurons is selective for the 3D orientation of texture gradients. A partially overlapping population of neurons in the same region is selective for the 3D orientation of disparity gradients. The fact that the preferences for tilt defined by texture was highly correlated with that for tilt defined by disparity is a compelling demonstration of the coding of texture-defined tilt by IT neurons. The latter demonstration, using single-unit recording methods, supports the suggestion that the IT activation by texture-defined 3D shapes reported in a functional imaging study (Serenio et al., 2002) reflects texture-based 3D shape selectivity of IT neurons. Our results also imply that some of these neurons carry a relatively abstract representation of tilt and, most likely, 3D shape. Indeed, the tilt preference of these IT neurons generalizes over texture type and, even more importantly, over depth cue. This cue invariance for 3D shape preference can be viewed as the 3D complement of the previously reported cue invariance for 2D shape selectivity in IT neurons (Sary et al., 1993; Tanaka et al., 2001).

IT neurons are known to be selective to texture in addition to shape and color (Desimone et al., 1984; Komatsu and Ideura, 1993). The present study shows that in some IT neurons, part of this texture selectivity is related to coding of 3D shape and not merely to 2D texture per se. This conclusion is supported by the strong correlation between the preferred tilts defined by each of the two cues and by the greater tilt selectivity for monocular compared with binocular pre-

sensation of the texture patterns. The latter fits the more salient 3D perception for monocular compared with binocular presentations and is not predicted if the texture selectivity merely reflects 2D texture coding. The response of many of the neurons was also modulated by the texture type. The latter probably reflects a coding of texture as a material property of an object. The high correlations of the preferred tilts for the different texture types suggest a separable coding of 3D- and material-related texture cues, akin to the separable coding of shape and color.

In the present study, approximately half of the IT neurons were selective for only one of the two depth cues, indicating that even in a high-level area such as IT, a sizeable proportion of neurons are not selective for all depth cues. Similar observations have been made regarding the cue invariance of 2D shape selectivity in IT. Indeed, Vogels and Orban (1996) compared the selectivity of IT neurons for shapes defined by either luminance or relative motion and found that 45% of the shape-selective neurons were selective for only one of the two cues. Similarly, Tanaka et al. (2001) reported that 56% of their 2D shape-selective neurons were selective for either the disparity or the luminance cue but not for both.

The preferred tilt for the texture cue was mostly independent of slant, which indicates separable coding of slant and tilt. This is similar to what has been reported in area MT (Nguyenkim and DeAngelis, 2003) for surfaces defined by disparity. We found that more neurons were significantly tuned for tilt than for slant. This might be because of the restricted range of slants (30°) that we used, which was necessary to avoid large changes in overall stimulus size. In contrast, it is possible that, as with MT neurons (Nguyenkim and DeAngelis, 2003), IT neurons are indeed more sensitive to tilt than to slant.

In IT, disparity-defined 3D shape-selective neurons are found predominantly in TEs, part of the lower bank of the STS (Janssen et al., 2000b), which is where we recorded in the present study. Because our strategy for determining texture-defined 3D selectivity requires disparity-defined 3D selectivity, it cannot be applied to neurons in other parts of IT not possessing disparity-based 3D shape selectivity. Other manipulations will be required to assess whether or not texture-based 3D shape selectivity is confined to a restricted part of IT, as seems to be the case for stereo-based 3D shape selectivity.

The degree of tilt tuning in this IT region appears to be similar to that found in the parietal area CIP (Tsutsui et al., 2002) but stronger than the tilt modulations observed in MT for disparity gradients (Nguyenkim and DeAngelis, 2003). Although comparisons between studies using somewhat different stimuli and search protocols require caution, it is tempting to conclude that CIP and IT are more intimately involved in the coding of higher order depth relationships than area MT. The present results also show that convergence of the stereo and texture cues for surface tilt is at least as strong in IT as in CIP. At present, it is unclear what the respective contributions of dorsal area CIP and ventral area IT are for 3D shape coding. Neurons in the two areas are selective for higher order disparity and texture depth cues, although selectivity for curvature still needs to be demonstrated in CIP. In both areas, disparity selectivity is abolished for anticorrelated random dot stereograms, as it is in perception (Naganuma et al., 2002; Janssen et al., 2003). It is possible that CIP codes the depth structure of surfaces, and that this information is then further processed into object descriptions in parietal area AIP and in IT for grasping (Murata et al., 2000) and recognition of objects, respectively. Alternatively, the depth information in IT might be the result of input from ventral rather than dorsal visual areas. In-

deed, it has been shown that V4 neurons are selective for the 3D orientation of bars (Hinkle and Connor, 2002), but it is unclear whether these neurons are sensitive to surface disparities, as IT neurons, or to orientation disparities.

Furthermore, it is unknown whether V4 neurons show convergence of multiple depth cues. Future research is necessary to determine whether the 3D cue invariance is created within IT or derived from 3D cue invariant input from other ventral or dorsal areas.

## References

- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051–2062.
- Gibson JJ (1950) *The perception of the visual world*. Boston: Houghton Mifflin.
- Hinkle DA, Connor CE (2002) Three-dimensional orientation tuning in macaque area V4. *Nat Neurosci* 5:665–670.
- Howard IP, Rogers BJ (2002) Seeing in depth. In: *Depth perception*, Vol 2. Toronto: Porteous.
- Janssen P, Vogels R, Orban GA (1999) Macaque inferior temporal neurons are selective for disparity-defined 3D shape. *Proc Natl Acad Sci USA* 96:8217–8222.
- Janssen P, Vogels R, Orban GA (2000a) Three-dimensional shapes coding in inferior temporal cortex. *Neuron* 27:385–397.
- Janssen P, Vogels R, Orban GA (2000b) Selectivity for three-dimensional shape that reveals distinct areas in macaques inferior temporal cortex. *Science* 288:2054–2056.
- Janssen P, Vogels R, Liu Y, Orban GA (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *J Neurosci* 21:9419–9429.
- Janssen P, Vogels R, Liu Y, Orban GA (2003) At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron* 37:693–701.
- Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20:535–538.
- Komatsu H, Ideura Y (1993) Relationships between color, shape, and pattern selectivities of neurons in the inferior temporal cortex of the monkey. *J Neurophysiol* 70:677–694.
- Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol* 83:2580–2601.
- Naganuma T, Tsutsui K, Sakata H, Taira M (2002) Surface-orientation-selective neurons in parietal area CIP and perception of 3D surface orientation. *Soc Neurosci Abstracts* 56:15.
- Nguyenkim JD, DeAngelis GC (2003) Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J Neurosci* 23:7117–7128.
- Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995–997.
- Sereno ME, Trinath T, Augath M, Logothetis NK (2002) Three-dimensional shape representation in monkey cortex. *Neuron* 33:635–652.
- Taira M, Tsutsui KJ, Jiang M, Yara K, Sakata H (2000) Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 83:3140–3146.
- Tanaka H, Uka T, Yoshiyama K, Kato M, Fujita I (2001) Processing of shape defined by disparity in monkey inferior temporal cortex. *J Neurophysiol* 85:735–744.
- Tanaka K, Saito H, Fukada Y, Morioka M (1991) Coding visual images of objects in the inferotemporal cortex of macaque monkey. *J Neurophysiol* 66:170–189.
- Tsutsui KI, Sakata H, Naganuma T, Taira M (2002) Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298:409–412.
- Vogels R, Orban GA (1994) Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *J Neurophysiol* 71:1428–1451.
- Vogels R, Orban GA (1996) Coding of stimulus invariances by inferior temporal neurons. *Prog Brain Res* 112:195–211.